E. Charniak, G. Carroll, J. Adcock, A. Cassandra, Y. Gotoh, J. Katz, M. Littman and J. McCann, Taggers for parsers

We consider what tagging models are most appropriate as front ends for probabilistic context-free grammar parsers. In particular, we ask if using a "multiple tagger", a tagger that returns more than one tag, improves parsing performance. Our conclusion is somewhat surprising: single-tag Markov-model taggers are quite adequate for the task. First of all, parsing accuracy, as measured by the correct assignment of parts of speech to words, does not increase significantly when parsers select the tags themselves. In addition, the work required to parse a sentence goes up with increasing tag ambiguity, though not as much as one might expect. Thus, for the moment, single taggers are the best taggers.

A. Srinivasan, S.H. Muggleton, M.J.E. Sternberg and R.D. King, Theories for mutagenicity: a study in first-order and feature-based induction

A classic problem from chemistry is used to test a conjecture that in domains for which data are most naturally represented by graphs, theories constructed with inductive logic programming (ILP) will significantly outperform those using simpler feature-based methods. One area that has long been associated with graph-based or structural representation and reasoning is organic chemistry. In this field, we consider the problem of predicting the mutagenic activity of small molecules: a property that is related to carcinogenicity, and an important consideration in developing less hazardous drugs. By providing an ILP system with progressively more structural information concerning the molecules, we compare the predictive power of the logical theories constructed against benchmarks set by regression, neural, and tree-based methods.

M. Pradhan, M. Henrion, G. Provan, B. del Favero and K. Huang, The sensitivity of belief networks to imprecise probabilities: an experimental investigation

Bayesian belief networks are being increasingly used as a knowledge representation for reasoning under uncertainty. Some researchers have questioned the practicality of obtaining the numerical probabilities with sufficient precision to create belief networks for large-scale applications. In this work, we investigate how precise the probabilities need to be by measuring how imprecision in the probabilities affects diagnostic performance. We conducted a series of experiments on a set of real-world belief networks for medical diagnosis in liver and bile disease. We examined the effects on diagnostic performance of (1) varying the mappings from qualitative frequency weights into numerical probabilities, (2) adding random noise to the numerical probabilities, (3) simplifying from quaternary domains for diseases and findings—absent, mild, moderate, and severe—to binary domains—absent and present, and (4) using test cases that contain diseases outside the network. We found that even extreme differences in the probability mappings and large amounts of noise lead to only modest reductions in diagnostic performance. We found no significant effect of the simplification from quaternary to binary representation. We also found that outside diseases degraded performance modestly. Overall, these findings indicate that even highly imprecise input probabilities may not impair diagnostic performance significantly, and that simple binary representations may often be adequate. These findings of robustness suggest that belief networks are a practical representation without requiring undue precision.

D. Whitney, S. Rana, J. Dzubera and K.E. Mathias, Evaluating evolutionary algorithms

Test functions are commonly used to evaluate the effectiveness of different search algorithms. However, the results of evaluation are as dependent on the test problems as they are on the algorithms that are the subject of comparison. Unfortunately, developing a test suite for evaluating competing search algorithms is difficult